Transition choice probabilities and welfare in ARUM's

André de Palma, Karim Kilani

ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

TRANSITION CHOICE PROBABILITIES AND WELFARE IN ARUM'S

André de PALMA
Karim KILANI

June 2009

Cahier n° 2009-32

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.enseignement.polytechnique.fr/economie/
mailto:chantal.poujouly@polytechnique.edu

Transition choice probabilities and welfare in ARUM's

André de Palma[1] · **Karim Kilani**[2]

June 2009 / First draft: January 2006

Abstract We study the descriptive and the normative consequences of price and/or other attributes changes in additive random utility models. We first derive expressions for the transition choice probabilities associated to these changes. A closed-form formula is obtained for the logit. We then use these expressions to compute the cumulative distribution functions of the compensating variation conditional on ex-ante and/or ex-post choices. The unconditional distribution is also provided. The conditional moments of the compensating variation are obtained as a one-dimensional integral of the transition choice probabilities. This framework allows us to derive a stochastic version of Shephard's lemma, which relates the expected conditional compensating variation and the transition choice probabilities. We compute the compensating variation for a simple binary linear in income choice model and show that the information on the transitions leads to better estimates of the compensating variation than those obtained when only ex-ante or ex-post information on individual choices is observed. For the additive in income logit, we compute the conditional distribution of compensating variation, which generalizes the logsum formula. Finally, we derive a new welfare formula for the disaggregated version of the representative consumer CES model.

Keywords Additive random utility models (ARUM) · Logit · Transition choice probabilities · Compensating variation · Shephard's Lemma · Logsum · CES

JEL classification D11 · D60

[1] Professor, Ecole Normale Supérieure de Cachan, Cachan 94230, France, and Ecole Polytechnique, France. E-mail: andre.depalma@ens-cachan.fr. Tel: +33-1-4740-5575. Fax: +33-1-4640-2460.

[2] Associate-Professor, Conservatoire National des Arts et Métiers, Paris 75003, France

1 Introduction

Discrete choice models (DCM) describe the individual choices of one alternative among a set of mutually exclusive alternatives. In the standard approach, each alternative i is associated with a utility $U_i = v_i + \varepsilon_i$, where v_i containing price and/or other attributes is the systematic utility and ε_i is an error term known by the individual but treated as a random variable by the modeler. The individual selects the alternative with the largest utility. The modeler assigns a probability $\mathbb{P}_i$ that an individual selects alternative i, which is the probability that the random variable U_i is larger than all the other random variables U_k, $k \neq i$. This approach corresponds to the random utility maximization models (RUM). Here, the systematic utility is fixed and the individual choices are static. Conversely, the rationalization and consistency of choice probabilities with random utility maximization probabilities is a well-known problem which is discussed e.g. in McFadden (2005).

Such models have initially been studied in the transport literature (to describe the choice between private and public transportation) and in the urban literature (to describe residential location; see the early contributions of Domencich and McFadden 1975). Later on, RUM models have been used in many other fields, such as education, demography, industrial organization, public economics, experimental economics, decision theory and marketing (see Anderson, de Palma and Thisse 1992, who have discussed the neoclassical economic foundations of RUM; see also the survey of McFadden 2001). Estimation of RUM (logit, probit, ordered probit, generalized extreme value models, mixed logit, etc.) has attracted a lot of attention during the last half century (see, e.g. Train 2003).

The welfare properties of RUM's are well known for the simplest model, the multinomial logit, which leads to the "logsum" formula (in the standard logit, ε_i are i.i.d. double-exponentially distributed and income enters the utility function linearly with a uniform coefficient, i.e. there are no income effects). The extensions including income effects and using other error terms specifications are more intricate.

Small and Rosen (1981) have addressed the question of income effects in RUM. They have derived an approximative expression of the expected compensating variation for a price or other attributes change (they focus on taxation). They extend the conventional welfare approach to the DCM framework and show that the expected compensating variation can be computed as an integral of the Hicksian choice probabilities (compensated choice probabilities). Using a similar approach, Dadgsvik and Karlström (2005) derive an exact formula for the compensating variation (CV) associated to a price (or other attributes) changes. They provide an expression for the distribution of the CV conditional on the ex-ante (i.e. before the change) individual choices, i.e. given that the individual choices are only observed ex-ante.

Welfare measures with income effects have also been studied via numerical simulations by McFadden (1999), who has developed a sampler for computing the CV caused by a change in the individual environment. For the generalized extreme value (GEV) models, which extends the multinomial logit model, he has provided an algorithm, the GEV sampler, to estimate welfare effects. However, even though this sampler leads to consistent results, it is time consuming since a large number of iterations must be performed in order to obtain with a reasonable level of accuracy numerical approximations of the true welfare impacts.

Von Haefen (2003) has worked out an application which suggests that the observed choice behavior of the individuals ex ante improves the accuracy of the calculation of consumer surplus. The scope of this paper is to analyze the theoretical properties of demand and welfare of RUM when a price or attribute change has occurred and when ex ante and/or ex post choices are observed.

As a consequence of price or attribute change, some individuals alter their ex ante choices. It is assumed that individual error terms remain the same ex-ante and ex-post (i.e. after the change). The expressions for the transition choice probabilities are derived analytically in this paper. This information is useful *per se* in order to evaluate the consequences of a policy but it is not sufficient to observe only the choice probabilities ex-ante and/or ex-post. Via an example, we show the information on transition choice probabilities is crucial to evaluate the welfare consequences of this change. We compute the transition choice probabilities and the distribution of the CV conditional on the transition choice. We use a direct approach based on Marshalian transition choice probabilities while Dagsvig and Karlström (2005) use Hicksian static choice probabilities relying on unobservable information, since their values depend on the unobservable error terms ε_i.

The structure of the paper is as follows. In Section 2, we provide the assumptions on the utility functions and on the distribution of the error terms. Theorem 1 provides an analytical formula for the transition choice probabilities for additive random utility maximization models (ARUM). The logit special case is handled in Proposition 2. In Section 3, we define the CV for ARUM. Theorem 4 provides an analytical expression (based on the transition choice probabilities) for the distribution of the CV conditional on the transitions. The various moments of the CV are given as a one-dimensional integral either of the transition choice probabilities (Theorem 6) or of the choice probabilities (Corollary 7). We also introduce a stochastic version of Shephard's Lemma for DCM (Proposition 8) in the context of transitions. In Section 4, we compute the CV for a simple binary linear in income RUM and consider the impacts of a change in one price. This example shows that the information on the transitions leads to better estimates of the CV than the ones obtained when only ex-ante or ex-post information on individual choices are observed. For the additive in income logit, we compute directly the conditional distribution of CV which generalizes the celebrated logsum formula. Finally, for the disaggregated version of the CES, we propose a new exact welfare measure. In Section 5 we discusses further extensions. Proofs are relegated to the appendix.

2 Transition choice probabilities

2.1 Model and notations

There are n alternatives and preferences are described by an ARUM. We consider the impacts of a change and study the individual choices before (ex-ante) and after (ex-post) the change. The ex-ante (conditional) utility U_i of an individual selecting i is given by $U_i = v_i + \varepsilon_i$, where v_i, the ex-ante systematic component of the utility U_i of i is assumed to be observable and where ε_i is an error term, which captures unobservable individual characteristics that are modelled by the econometric investigator as a random variable. We assume that the error terms remain the same ex-ante and ex-post.

Let F be the CDF of the vector of error terms $(\varepsilon_1 ... \varepsilon_n)$ which is assumed to be absolutely continuous with respect to the Lebesgue measure over a convex support. Therefore (see McFadden 1978) the probability $\mathbb{P}_i(\underline{v})$, that an individual selects ex-ante i can be written in an integral form (the function's arguments are omitted in the sequel when it is unambiguous)

$$\mathbb{P}_i(\underline{v}) \equiv \Pr(U_i > U_k, k \neq i) = \int_{-\infty}^{+\infty} F^i(u - v_1 ... u - v_n)\, du, \tag{1}$$

where $\underline{v} \equiv (v_1 ... v_n)$ is the systematic utility vector and where: $F^i(x_1 ... x_n) \equiv \partial F(x_1 ... x_n) / \partial x_i$. It can easily be verified that: $\sum_i \mathbb{P}_i = 1$. Note that the choice probabilities are invariant up to a shift: $\mathbb{P}_i(v_1 + \delta ... v_n + \delta) = \mathbb{P}_i(\underline{v})$. The expected individual demand $\mathbb{X}_i$ for alternative i can be obtained by using Roy's identity (see Section 4 for an illustration in the CES case).

Let Π_i^j be minus the derivative of $\mathbb{P}_i$ with respect to v_j. A derivation of (1) under the integral sign (see Anderson et al. 1992) yields:

$$\Pi_i^j \equiv -\frac{\partial \mathbb{P}_i}{\partial v_j} = \int_{-\infty}^{+\infty} F^{ij}(u - v_1 ... u - v_n)\, du, \tag{2}$$

where $F^{ij} \equiv \partial F^i / \partial x_j$, $i, j = 1 ... n$. Note the equality of the cross-derivatives: $\Pi_i^j = \Pi_j^i$, $j \neq i$.

The ex-post utility of an individual selecting j is $\Upsilon_j = \omega_j + \varepsilon_j$, where ω_j is the (observable) ex-post systematic component of Υ_j. The probability of selecting ex-post j is given by $\mathbb{P}_j(\underline{\omega})$, where $\underline{\omega} \equiv (\omega_1 ... \omega_n)$ (see Eq. (1)).

2.2 Computation of transition probabilities

The (transition) choice probability that an individual selects i ex-ante and j ex-post is

$$\mathbb{P}_{i \hookrightarrow j}(\underline{v}; \underline{\omega}) \equiv \Pr(U_i > U_k, k \neq i; \Upsilon_j > \Upsilon_r, r \neq j). \tag{3}$$

Theorem 1 provides an integral form for these transition choice probabilities. Let $\delta_k \equiv \Upsilon_k - U_k = \omega_k - v_k$, $k = 1 ... n$, be the utility variation of k. We assume without loss of generality the ranking $\delta_1 \leq ... \leq \delta_n$. Define $t^+ = \max(t, 0)$. We have (see appendix):

Theorem 1 *For an ARUM, consider the change: $\underline{v} \to \underline{\omega}$. The transition choice probabilities from i to j are given by:*

$$\mathbb{P}_{i \hookrightarrow j}\left(\underline{v};\underline{\omega}\right) = \begin{cases} \mathbb{P}_i\left(v_1 + (\delta_1 - \delta_i)^+ ... v_n + (\delta_n - \delta_i)^+\right), & \text{if } j = i; \\ \int_{\delta_i}^{\delta_j} \Pi_i^j\left(v_1 + (\delta_1 - z)^+ ... v_n + (\delta_n - z)^+\right) dz, & \text{if } j > i; \\ 0, & \text{if } j < i. \end{cases} \quad (4)$$

The probability $\mathbb{P}_{i \hookrightarrow i}$ to select i ex-ante and ex-post is given by a choice probability as defined by (1). We discuss the case $1 < i < n$, with $n > 2$ (the other cases are left to the reader). For $k < i$, $\delta_k \leq \delta_i$; therefore, if an individual selects i (with utility v_i) ex-ante, he will prefer i to k ex-post. Let $k > i$ with $\delta_k \geq \delta_i$. In this case, an individual who selects i ex-post (with utility ω_i) prefers i to k ex-ante. Therefore,

$$\mathbb{P}_{i \hookrightarrow i} = \mathbb{P}_i(v_1 ... v_i, \omega_{i+1} - \delta_i ... \omega_n - \delta_i) = \mathbb{P}_i(v_1 + \delta_i ... v_{i-1} + \delta_i, \omega_i ... \omega_n) \quad (5)$$

represents the probability that an individual selects i ex-ante and ex-post.

The transition choice probabilities from i to j, $j \neq i$ are clearly zero if j is weakly deteriorated in relative term with respect to i ($\delta_j \leq \delta_i$). For $j \geq i$, we define the transition $i \hookrightarrow j$ to be *feasible* if it occurs with a strictly positive probability. The transition choice probabilities are explained intuitively below. For $\delta_j > \delta_i$, these transition choice probabilities $\mathbb{P}_{i \hookrightarrow j}$ are given by an integral on $z = (\omega_j + \varepsilon_j) - (v_i + \varepsilon_i)$, which represents the utility variation of an individual who shifts from i to j. Note that $z > \delta_i$ (the utility variation when staying in i) and $z < \delta_j$ (otherwise j would have been preferred to i to ex ante). The integrand $\Pi_i^j\left(v_1 + (\delta_1 - z)^+ ... v_n + (\delta_n - z)^+\right)$ represents the probability density that the individual who experienced a utility change of z shifts from i to j. Finally note that the argument z in $\left(v_1 + (\delta_1 - z)^+ ... v_n + (\delta_n - z)^+\right)$ plays a similar role than δ_i in the vector $\left(v_1 + (\delta_1 - \delta_i)^+ ... v_n + (\delta_n - \delta_i)^+\right)$.

When $n = 2$ or 3, transition choice probabilities reduce to choice probabilities (using standard constraints on probabilities). For $n > 3$, there are a priori $n(n-1)/2$ integrals. However, using the $2n - 3$ constraints, the computation of all transition choice probabilities requires the computation of at most $(n-2)(n-3)/2$ integrals.

The constraints on the transition choice probabilities can be easily checked. As expected, the ex-ante and ex-post choice probabilities can be recovered by summation of the transition choice probabilities given in Theorem 1. More precisely, using (4) it can be shown that:

$$\sum_j \mathbb{P}_{i \hookrightarrow j} = \mathbb{P}_i\left(\underline{v}\right) \text{ and} \sum_i \mathbb{P}_{i \hookrightarrow j} = \mathbb{P}_j\left(\underline{\omega}\right). \quad (6)$$

Note that these expressions are straightforward to derive if one uses directly the expressions in (3).

2.3 Logit transition probabilities

The transition choice probabilities are explicit for the logit model. In this case, the CDF is given by:

$$F(x_1 ... x_n) = \exp\left(-\sum_i e^{-x_i}\right), \tag{7}$$

which yields the following choice probabilities (see Domencich and McFadden 1975):

$$\mathbb{P}_i(\underline{v}) = \frac{e^{v_i}}{\sum_k e^{v_k}}. \tag{8}$$

We will use in the rest of the paper the following notations:

$$\begin{cases} s_r \equiv \sum_{k \leq r} e^{v_k}; \\ \sigma_r \equiv \sigma_0 - \sum_{k \leq r} e^{\omega_k}, \text{ with } \sigma_0 \equiv \sum_k e^{\omega_k}; \\ \Omega_r \equiv s_r + \sigma_r e^{-\delta_r}, \ r = 1 ... n. \end{cases} \tag{9}$$

We have (see appendix):

Proposition 2 *For the logit specification (7), consider the change: $\underline{v} \to \underline{\omega}$. The transition choice probabilities from i to j are given by:*

$$\mathbb{P}_{i \hookrightarrow j} = \begin{cases} \frac{e^{v_i}}{\Omega_i}, & \text{if } j = i; \\ \sum_{r=i}^{j-1} \left(\frac{e^{v_i}}{\Omega_{r+1}} - \frac{e^{v_i}}{\Omega_r}\right) \frac{e^{\omega_j}}{\sigma_r}, & \text{if } j > i; \\ 0, & \text{if } j < i. \end{cases} \tag{10}$$

First remark that in the case of Example 1 with $\omega_1 < v_1$, $\mathbb{P}_{1 \hookrightarrow j}$, with $j \neq 1$, can be written, in the logit case, as:

$$\mathbb{P}_{1 \hookrightarrow j} = \left[\frac{e^{v_1}}{\sum_k e^{v_k}} - \frac{e^{\omega_1}}{e^{\omega_1} + \sum_{k>1} e^{v_k}}\right] \times \frac{e^{v_j}}{\sum_{k>1} e^{v_k}},$$

where the first term on the RHS represents the probability that an individual abandons 1, while the second term is the probability that j is the second best choice (this independence results is specific to the logit). The other cases are more involved and explained below.

Note that e^{v_i}/Ω_r, $r \geq i$ represent the probability to choose i ex-ante and to get a utility variation in $[\delta_i, \delta_r]$. The probability of this event can be written as $\Pr\left(U_i > U_k + (\delta_k - \delta_r)^+, k \neq i\right)$; it corresponds to a choice probability with the systematic utility given by $(v_1 ... v_r, \omega_{r+1} - \delta_r ... \omega_n - \delta_r)$. In particular, if $r = i$, e^{v_i}/Ω_i is the probability to have a utility variation of exactly δ_i. It corresponds to $\mathbb{P}_{i \hookrightarrow i}$ since the individual sticks to alternative i iff he has a utility variation of δ_i. Note that $\sum_{r=i}^{j-1} [(e^{v_i}/\Omega_{r+1} - e^{v_i}/\Omega_r)] = e^{v_i}/\Omega_j - e^{v_i}/\Omega_i$ represents the probability that an individual chooses i ex-ante and incurs a utility variation in $[\delta_i, \delta_j]$. If the individual shifts from i to j, the associated utility variation lies within the interval $[\delta_i, \delta_j]$. The term $e^{v_i}/\Omega_{r+1} - e^{v_i}/\Omega_r$ represents the probability that an individual abandon i and has a utility variation in the interval $[\delta_r, \delta_{r+1}]$. He will choose an alternative k such that $k > r$. The probability that he chooses j among the feasible choices k (with $k > r$) is $e^{\omega_j} / \sum_{k>r} e^{\omega_k}$. The reader is also referred to de Palma and Kilani (2005) who compute the conditional transition probabilities, where changes are conditional to the ex-ante choice.

3 Welfare

In the previous section, we provided an expression for the transition choice probabilities $\mathbb{P}_{i \hookrightarrow j}$ for a change $\underline{v} \to \underline{\omega}$. We study now the distribution of individual compensations and the welfare impacts associated to this change. We assume that the ex-ante (ex-post) indirect utility U_k (resp. Υ_k) of k is a function of the individual's income y. They are denoted as $U_k(y)$ (resp. as $\Upsilon_k(y)$) and assumed to be strictly increasing and continuous in y.

3.1 Welfare distributions and moments

The compensating variation cv is defined as the amount of income needed to restore the ex-ante individual's utility level ex-post change $\underline{v} \to \underline{\omega}$. In the DCM literature (see, McFadden 1999), this means:

$$\max_k (U_k) = \max_k \left[\Upsilon_k (y - cv)\right]. \tag{11}$$

Since the utilities are random due to the presence of the error terms (recall $U_i = v_i + \varepsilon_i$), cv is also a random variable.

In order to insure that Eq. (11) admits a unique solution, we should make an additional assumption. Let $\delta_k(c) \equiv \Upsilon_k(y - c) - U_k$ be the (deterministic) utility variation of k ex-post and after compensation of $-c$, with $\delta_k(0) = \delta_k$. We require that for any i, k, there exists a real ψ_{ik} defined by:

$$\delta_k(\psi_{ik}) = (\delta_k - \delta_i)^+ . \tag{12}$$

The interpretation of the $(\psi_{ik})' s$ is provided in the following Lemma (see appendix):

Lemma 3 *Given a feasible transition $i \hookrightarrow j$, the support of cv is included in $[m_{ij}, \overline{m}_j]$, where $m_{ij} \equiv \max\left(\psi_{ii}, \psi_{ij}\right)$ and where $\overline{m}_j \equiv \max_k \left(\psi_{jk}\right)$.*

As we have seen in Section 4.1, the CV conditional on the transitions $i \hookrightarrow i$ can be stochastic. This is not the case in the absence of income effects.

We wish to compute the distribution of cv using the information on the individual transitions ex-post: $\underline{v} \to \underline{\omega}$. Consider a feasible transition $i \hookrightarrow j$.

The CDF of cv, conditional on a feasible transition $i \hookrightarrow j$, denoted by $\Phi_{i\hookrightarrow j}$, is given by:

$$\Phi_{i\hookrightarrow j}(c) \equiv \frac{\Pr\left(c \geq cv; U_i > U_k, k \neq i; \Upsilon_j > \Upsilon_r, r \neq j\right)}{\mathbb{P}_{i\hookrightarrow j}\left(\underline{v}, \underline{\omega}\right)}. \tag{13}$$

In Theorem 4, an analytic expression for $\Phi_{i\hookrightarrow j}$ is provided. Let $\delta_k^+(c) = \max\left(\delta_k(c), 0\right)$ and recall that $m_{ij} \equiv \max\left(\psi_{ii}, \psi_{ij}\right)$ and $\overline{m}_j \equiv \max_k\left(\psi_{jk}\right)$. We have (see appendix):

Theorem 4 *For an ARUM, consider the change: $\underline{v} \to \underline{\omega}$. The CDF of the compensating variation conditional on the transition $i \hookrightarrow j$ has support $(m_{ij}, \overline{m}_j]$ and is given by:*

$$\Phi_{i\hookrightarrow j}(c) = \frac{\mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) \ldots v_n + \delta_n^+(c); \underline{\omega}\right)}{\mathbb{P}_{i\hookrightarrow j}\left(\underline{v}, \underline{\omega}\right)}, \ c \geq m_{ij}, \tag{14}$$

where the transition choice probabilities $\mathbb{P}_{i\hookrightarrow j}(.,.)$ are given in Theorem 1.

This expression allows the computation of the distribution of cv when only the ex-ante or the ex-post choice is observed. In this case, the conditional distribution of cv depends on the choice probabilities and not on the transition choice probabilities as in Theorem 4. We now compute $\Phi_{i\hookrightarrow}$ (resp. $\Phi_{\hookrightarrow j}$) the conditional CDF of cv given the ex-ante (resp. ex-post) choice of i (resp. j). Let $\underline{m}_j \equiv \min_i\left(m_{ij}\right)$ and let $H_{m_{ij}}(c) \equiv 1$ if $c \geq m_{ij}$ and $H_{m_{ij}}(c) \equiv 0$ otherwise be the Heaviside function at m_{ij}. We obtain (see appendix):

Corollary 5 *For an ARUM, consider the change: $\underline{v} \to \underline{\omega}$. The CDF of the compensating variation*
(a) conditional on the ex-ante choice of i has support $[\psi_{ii}, \overline{m}_n]$ and is:

$$\Phi_{i\hookrightarrow}(c) = \frac{\mathbb{P}_i\left(v_1 + \delta_1^+(c) \ldots v_n + \delta_n^+(c)\right)}{\mathbb{P}_i\left(\underline{v}\right)}, \ c \geq \psi_{ii}; \tag{15}$$

(b) conditional on the ex-post choice of j, has support $\left[\underline{m}_j, \overline{m}_j\right]$ and is:

$$\Phi_{\hookrightarrow j}(c) = \frac{\sum_i H_{m_{ij}}(c) \times \mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) \ldots v_n + \delta_n^+(c); \underline{\omega}\right)}{\mathbb{P}_j\left(\underline{\omega}\right)}, \ c \geq \underline{m}_j. \tag{16}$$

The CDF (15) coincides with the CDF derived by Dagsvik and Karlström (2005) in the case where only the ex-ante choices are observed. Note that for the logit model, the CDF of the CV conditional on the ex-ante choice of i is given by:

$$\Phi_{i\hookrightarrow}(c) = \frac{\sum_k e^{v_k}}{\sum_k e^{v_k + \delta_k^+(c)}}, \ c \geq \psi_{ii}. \tag{17}$$

Finally, the unconditional distribution of cv can be computed using Eq. (15) and making use of the theorem on total probability (see also Dagsvik and Karlström 2005):

$$\Phi(c) = \sum_i H_{m_{ii}}(c) \times \mathbb{P}_i\left(v_1 + \delta_1^+(c) \ldots v_n + \delta_n^+(c)\right).$$

We now compute the conditional and the unconditional moments of the distribution of cv (see appendix):

Theorem 6 *For an ARUM, consider the change: $\underline{v} \to \underline{\omega}$. The pth moment ($p \geq 1$) of the compensating variation conditional on the transition $i \hookrightarrow j$ is given by:*

$$\mathbb{E}_{i\hookrightarrow j}\left[cv^p\right] = \overline{m}_j^p - p\int_{m_{ij}}^{\overline{m}_j} c^{p-1}\frac{\mathbb{P}_{i\hookrightarrow j}\left(v_1+\delta_1^+(c)\,...v_n+\delta_n^+(c)\,;\underline{\omega}\right)}{\mathbb{P}_{i\hookrightarrow j}\left(\underline{v},\underline{\omega}\right)}dc. \tag{18}$$

When $p = 1$, Eq. (18) provides the expected CV conditional on the observed transitions. This is reminiscent of the standard treatment of surplus, and involves the computation of areas under the compensated transition choice probabilities curves. The conditional on the ex-ante or ex-post choices moments are given by (see appendix):

Corollary 7 *For an ARUM, consider the change: $\underline{v} \to \underline{\omega}$. The pth ($p \geq 1$) moment of the compensating variation conditional is given for*
(a) the ex-ante choice of i by:

$$\mathbb{E}_{i\hookrightarrow}\left[cv^p\right] = \overline{m}_n^p - p\int_{\psi_{ii}}^{\overline{m}_n} c^{p-1}\frac{\mathbb{P}_i\left(v_1+\delta_1^+(c)\,...v_n+\delta_n^+(c)\right)}{\mathbb{P}_i\left(\underline{v}\right)}dc; \tag{19}$$

(b) the ex-post choice of j by:

$$\mathbb{E}_{\hookrightarrow j}\left[cv^p\right] = \overline{m}_j^p - p\sum_i\int_{\psi_{ii}}^{\overline{m}_j} c^{p-1}\frac{\mathbb{P}_{i\hookrightarrow j}\left(v_1+\delta_1^+(c)\,...v_n+\delta_n^+(c)\,;\underline{\omega}\right)}{\mathbb{P}_j\left(\underline{\omega}\right)}dc. \tag{20}$$

Equation (19) provides in particular the expected CV conditional on the ex-ante choice (this expression is derived in Dagsvik and Karlström 2005). It involves the computation of areas under the compensated choice probability curves. Equation (20) is new and relies on the expression obtained in Theorem 6.

Using Corollary 7 with Eq. (6), the pth unconditional moment of the CV verifies:

$$\mathbb{E}\left[cv^p\right] = \overline{m}_n^p - p\sum_i\int_{\psi_{ii}}^{\overline{m}_n} c^{p-1}\mathbb{P}_i\left(v_1+\delta_1^+(c)\,...v_n+\delta_n^+(c)\right)dc. \tag{21}$$

In particular, the expectation of cv is given by

$$\mathbb{E}\left[cv\right] = \overline{m}_n - \sum_i\int_{\psi_{ii}}^{\overline{m}_n} \mathbb{P}_i\left(v_1+\delta_1^+(c)\,...v_n+\delta_n^+(c)\right)dc. \tag{22}$$

According to Eq. (22), $\mathbb{E}\left[cv\right]$ is the sum of the integrals of parametrized choice probabilities $\mathbb{P}_i\left(v_1+\delta_1^+(c)\,...v_n+\delta_n^+(c)\right)$. An approximative expression for the expected CV was also envisaged by Small and Rosen (1981).

3.2 Shephard's lemma revisited

We assume that the systematic component of the utility (ex-ante and ex-post) of k depends on income y and on price level p_k and is given by $V_k(y, p_k)$. Assuming that $V_k(.,.)$ is differentiable with respect to both arguments, the conditional (individual) demand x_k for good k is determined by using Roy's identity:

$$x_k = -\frac{(\partial V_k/\partial p_k)}{(\partial V_k/\partial y)},\ k = 1...n.$$

Note that in ARUM, the conditional demands are deterministic, i.e. are independent on the idiosyncratic taste parameters. Let Δp_k be a price change of good k. The corresponding CV for an individual who sticks to good k is ψ_{kk}. Shephard's Lemma, which is a direct application of the Envelope Theorem, gives:

$$\lim_{\Delta p_k \to 0} \frac{\psi_{kk}}{\Delta p_k} = -x_k.$$

In the RUM approach, when an individual modify her choice after an infinitesimal price change, the corresponding CV is stochastic (i.e. depends on the idiosyncratic terms of the initial and of the final good). Therefore, we compute the expected CV, conditional of the transition in order to write the counterpart of Shephard's Lemma in the RUM models. We have (se appendix):

Proposition 8 *For an ARUM, consider the infinitesimal change of the price of one good. The expected change in CV per dollar for an infinitesimal price increase of good* 1*, conditional on the ex-ante and the ex-post choices is:*

$$\lim_{\Delta p_1 \to 0^+} \frac{\mathbb{E}_{1\hookrightarrow j}[cv]}{\Delta p_1} = \begin{cases} -x_1, & \textit{if } j = 1; \\ -\frac{\rho_{1j}}{2} x_1 & \textit{if } j > 1, \quad \rho_{1j} \leq 1; \\ -\left(1 - \frac{\rho_{j1}}{2}\right) x_1 & \textit{if } j > 1, \quad \rho_{j1} \leq 1, \end{cases} \tag{23}$$

where $\rho_{ij} \equiv (\partial V_i/\partial y) / (\partial V_j/\partial y)$.
The expected change in CV per dollar for an infinitesimal price decrease of good n*, conditional on the ex-ante and the ex-post choices is:*

$$\lim_{\Delta p_n \to 0^-} \frac{\mathbb{E}_{i\hookrightarrow n}[cv]}{\Delta p_n} = \begin{cases} -x_n/2, & \textit{if } i < n; \\ -x_n, & \textit{if } i = n. \end{cases} \tag{24}$$

Consider a price increase. The result for the case if $j = 1$ is trivial, since this is Sheppard's Lemma. The intuition for the case if $j > 1$ is as follows. First note that the consumer who are indifferent between 1 and j (i.e. the first individual to shift) requires no compensation. Second, consider the "last" individual ready to shift from 1 to j. She is indifferent between state 1 and state j. The indifference ex-post implies that: $v_1(p_1 + \Delta p_1, y) + \varepsilon_1 = v_j(p_j, y) + \varepsilon_j$. Since $\Delta p_1 \to 0$, we have: $\varepsilon_j - \varepsilon_1 = v_1 + \Delta p_1 (\partial V_1/\partial p_1) - v_j$ (where argument are omitted when unnecessary). The CV gives:

$$v_1(p_1, y) + \varepsilon_1 = v_j(p_j, y - cv) + \varepsilon_j.$$

Since $cv \to 0$ as $\Delta p_1 \to 0$, $v_1 + \varepsilon_1 = v_j - cv\,(\partial V_j/\partial y) + \varepsilon_j$, so that, using the expression for $\varepsilon_j - \varepsilon_1$ derived above, we get:

$$cv = \frac{v_j - v_1 + (\varepsilon_j - \varepsilon_1)}{\partial V_j/\partial y} = \Delta p_1 \frac{\partial V_1/\partial p_1}{\partial V_j/\partial y}.$$

Using Roy's identity $((\partial V_1/\partial p_1)\,/\,(\partial V_1/\partial y) = -x_1)$, we get, as required, that the average (per dollar) CV is $-x_1\rho_{1j}/2$.

Finally, note that by applying the theorem on total probability to (23) and (24), one obtains: $\lim_{\Delta p_1 \to 0^+} \mathbb{E}\,[cv]\,/\Delta p_1 = \mathbb{X}_1$ and $\lim_{\Delta p_n \to 0^-} \mathbb{E}\,[cv]\,/\Delta p_n = \mathbb{X}_n$, respectively. This weaker version of the Shephard's has been obtained by Dagsvik and Kalstrom (2005).

4 Examples

4.1 Welfare estimates with transition information

Consider two alternatives, denoted by 1 and 2, and we study the consequences of a price change. We show that the econometric investigator can get much better estimates of the welfare impacts of this change, when information concerning ex-ante choice and ex-post choice are used.

Assume that the ex-ante utility of a given individual is $U_i = \alpha_i\,(y - p_i) + \varepsilon_i$, where $\alpha_i > 0$ is the marginal utility of income (denoted by y) of good i, p_i is the prices of good i, and ε_i is an unobservable error term, $i = 1, 2$. Let $\eta \equiv U_1 - U_2 = \varepsilon_1 - \varepsilon_2$ uniformly distributed over $[-1, 1]$. Hence, good 1 is chosen ex-ante iff $\eta > 0$. We study the transition when the price of good 1 is raised by $\Delta p_1 > 0$, and we assume: $1 > \max\,(\alpha_1, \alpha_2)\,\Delta p_1$.

Three cases arise: (a) if $\eta > \alpha_1 \Delta p_1$, the individual chooses 1 ex-ante and ex-post; (b) if $\alpha_1 \Delta p_1 \geq \eta > 0$, the individual chooses 1 ex-ante and 2 ex-post; (c) if $\eta < 0$, the individual chooses 2 ex-ante and ex-post. The cv corresponding to this price change is given by: $cv = 0$ if $\eta \leq 0$; $cv = -\eta/\alpha_2$ if $0 < \eta \leq \alpha_2 \Delta p_1$; $cv = -\Delta p_1$ if $\alpha_2 \Delta p_1 < \eta$.

Let $\alpha_1 \geq \alpha_2$. Three cases arise: (a) For a transition $1 \hookrightarrow 1$, we have $cv = -\Delta p_1$ (b) For a transition $1 \hookrightarrow 2$, the support of cv is $[-\Delta p_1, 0]$. There is a mass at $(-\Delta p_1)$ corresponding to the probability that the individual shifts from 1 to 2, and returns to 1 after being compensated by $-cv$. Otherwise, the individual selects good 2 after being compensated by η/α_2. (c) For a transition

$2 \hookrightarrow 2$, $cv = 0$. The discussion is illustrated in Figure 1.

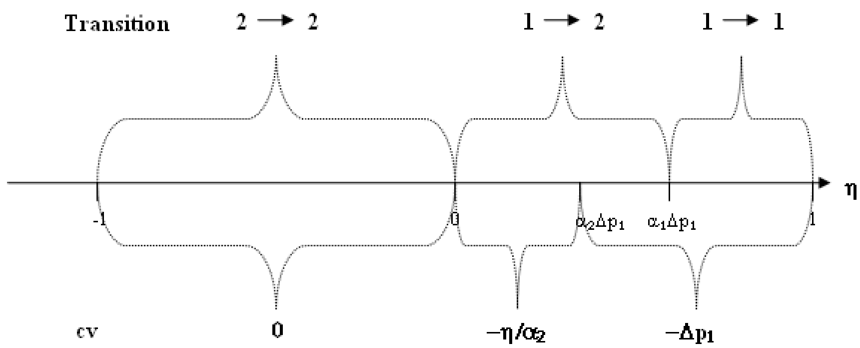

Figure 1: Transitions and CV with respect to η (case $\alpha_1 > \alpha_2$)

The case $\alpha_1 < \alpha_2$ is similar and left to the reader.

We can use the above discussion to compute the expected CV conditional to the transition $i \hookrightarrow j$, $i, j = 1, 2$, in the case: $\alpha_1 = \alpha_2 = \alpha$ (no income effects). We have: $\mathbb{E}_{1\hookrightarrow 1}(cv) = -\Delta p_1$; $\mathbb{E}_{1\hookrightarrow 2}(cv) = -\Delta p_1/2$; $\mathbb{E}_{2\hookrightarrow 2}(cv) = 0$.

We wish to compare the quality of the estimates of cv with respect to the knowledge of the ex-ante and/or ex-post choice. Without ex-ante and/or ex-post information concerning individual's choice, an appropriate estimate of cv is the expected CV: $\mathbb{E}(cv) = -(1 - \alpha\Delta p_1/2)(\Delta p_1/2)$.

First, assume that only the ex-ante choice is observed. If the individual selects 2 ex-ante, cv is deterministic and equal to 0, so that the conditional expectation denoted by $\mathrm{E}_{2\hookrightarrow}(cv)$ verifies: $\mathrm{E}_{2\hookrightarrow}(cv) = 0$. If the individual selects 1 ex-ante, cv is random and replaced by its conditional expectation given by: $\mathrm{E}_{1\hookrightarrow}(cv) = -(1 - \alpha\Delta p_1/2)\Delta p_1$.

Second, assume that only the ex-post choice is observed. If the individual selects 1 ex-post: $\mathbb{E}_{\hookrightarrow 1}(cv) = -\Delta p_1$. If the individual selects 1 ex-post, we get: $\mathbb{E}_{\hookrightarrow 2}(cv) = -[\alpha\Delta p_1/(\alpha\Delta p_1 + 1)](\Delta p_1/2)$.

Third, assume that the ex-ante and the ex-post choices are observed. If 1 is selected ex-ante and ex-post, then $cv = -\Delta p_1$; if 2 is selected ex-ante and ex-post, then $cv = 0$. If 1 is selected ex-ante and 2 is selected ex-post then cv is random and replaced by its conditional expectation: $\mathbb{E}_{1\hookrightarrow 2}(cv) = -\Delta p_1/2$.

In summary: the individual in 2 ex-ante or in 1 ex-post receive a deterministic compensation. By contrast, the observation of the choice of 1 ex-ante *only* or of 2 ex-post *only* is insufficient: information on ex-ante and ex-post choices $(1 \hookrightarrow 2)$ improves the quality of information on the CV.

We have computed the root-mean square errors (R.M.S.E.) $\sigma(cv\,|\mathcal{I})$ for the four estimators based on the information $\mathcal{I}$ on individual choice: "without" information, with "ex-ante", with "ex-post" and with "transitions" information.

The largest gains occur when transitions are observed. When only ex-post information is available, the gain can be small. Figure 2 shows the impact of the magnitude of the change Δp_1 for $\alpha = 1$.

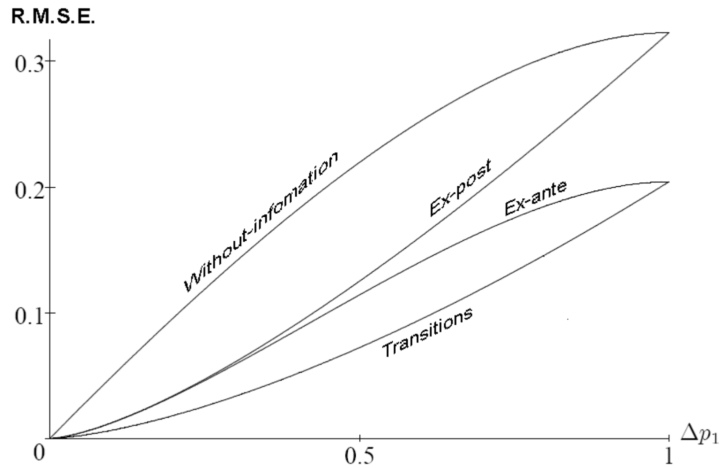

Figure 2: R.M.S.E. for various information regimes

These results suggest that the information on the ex-ante and/or ex-post individual choices lead to better estimates of the CV, but that an ex-ante information only is better than ex-post information only. When $\Delta p_1 = 1$, there are no more transitions so that "ex-ante" and "transitions" information regimes coincide. Similarly, "without" and with "ex-post" information regimes also coincide.

4.2 Additive in income logit specification

In this section, we concentrate our attention on the logit model where the transition choice probabilities have an explicit form (see Proposition 2). We assume that the utility is additive in income, i.e. that $U_k - v(y)$ (resp. $\Upsilon_k - v(y)$) is independent on income, where $v(.)$ is strictly increasing. In this case, the

We first provide the expressions for the CDF of cv conditional on the transition $i \hookrightarrow j$ which have closed forms (proof in appendix):

Proposition 9 *For the logit specification (7) with additive in income utility, consider the change: $\underline{v} \to \underline{\omega}$. The compensating variation conditional on the transition $i \hookrightarrow j$ has support $\left[\psi_{ii}, \psi_{jj}\right]$. For $c \in \left[\psi_{ll}, \psi_{(l+1)(l+1)}\right]$, $j > l \geq i$, the CDF is given by:*

$$\Phi_{i \hookrightarrow j}(c) = \frac{1}{\Xi_{ij}} \left[\Xi_{il} + \frac{1}{\sigma_l} \left(\frac{1}{s_l + \sigma_l e^{-\delta_y(c)}} - \frac{1}{\Omega_l} \right) \right], \tag{25}$$

where $\Xi_{ii} = 0$ and $\Xi_{il} = \sum_{r=i}^{l-1} \sigma_r^{-1} \left(\Omega_{r+1}^{-1} - \Omega_r^{-1} \right)$, $l > i$.

The expected CV conditional on the transition $i \hookrightarrow j$ can be computed up to $(n-1)$ integral terms (proof in appendix):

Proposition 10 *For the additive in income logit, consider the change: $\underline{v} \to \underline{\omega}$. The expected compensating variation conditional on the transition $i \hookrightarrow j$, $j > i$, is given by:*

$$\mathbb{E}_{i\hookrightarrow j}[cv] = \begin{cases} \psi_{ii}, & \text{if } j = i; \\ \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \frac{1}{\sigma_r} \left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}} - \frac{\psi_{rr}}{\Omega_r} - \theta_r \right], & \text{if } j > i, \end{cases} \tag{26}$$

where $\Xi_{ij} \equiv \sum_{r=i}^{j-1} \sigma_r^{-1} \left(\Omega_{r+1}^{-1} - \Omega_r^{-1} \right)$, $j > i$, and where

$$\theta_r \equiv \int_{\psi_{rr}}^{\psi_{(r+1)(r+1)}} \frac{dc}{s_r + \sigma_r e^{-\delta_y(c)}}, \quad r = 1...n-1, \tag{27}$$

with s_r, σ_r and Ω_r given by (9).

The formula (26) with (27) generalizes the standard logsum expression (discussed below) in many ways. It conditions the expected CV on both the ex-ante and the ex-post choices and it captures income effects.

Using the same integral terms θ_r $(r = 1...n-1)$, it is possible to derive expressions of the expected CV when the ex-ante or the ex-post (Corollary 11) are observed. We have (see appendix):

Corollary 11 *For the additive in income logit, consider the change: $\underline{v} \to \underline{\omega}$. The expected compensating variation conditional on*
(a) the ex-ante choice of i is:

$$\mathbb{E}_{i\hookrightarrow}[cv] = \begin{cases} \psi_{nn} - s_n \sum_{r=i}^{n-1} \theta_r, & \text{if } i < n; \\ \psi_{nn}, & \text{if } i = n; \end{cases} \tag{28}$$

(b) the ex-post choice of j is:

$$\mathbb{E}_{\hookrightarrow j}[cv] = \begin{cases} \psi_{11}, & \text{if } j = 1; \\ \sigma_0 \left\{ \frac{\psi_{jj}}{\sigma_{j-1}} - \sum_{r=1}^{j-1} \frac{1}{\sigma_r} \left(\frac{e^{\omega_r} \psi_{rr}}{\sigma_r} - s_r \theta_r \right) \right\}, & \text{if } j > 1, \end{cases} \tag{29}$$

with s_r, σ_r and Ω_r given by (9).

Given that $\mathbb{E}[cv] = \sum_{i=1}^{n} \mathbb{P}_i(\underline{v}) \mathbb{E}_{i\hookrightarrow}[cv]$, we get that for the additive in income logit, the expected CV is:

$$\mathbb{E}[cv] = \psi_{nn} - \sum_{r=1}^{n-1} s_r \theta_r. \tag{30}$$

Using for example Eq. (28), we get $\mathbb{E}[cv] = \psi_{nn} - \sum_{i=1}^{n-1} \sum_{r=i}^{n-1} e^{v_i} \theta_r$. Eq. (30) is obtained by inverting the two sum signs.

Assume for example that for all initial choice, the individual has benefited from the change. In this case, ψ_{nn} is the maximal benefit induced by this change. This benefit has to be reduced to take into account that the individual with another ex-ante choice requires a smaller compensation.

Proposition 10, Corollary 11 and Eq. (30) show that the conditional and the unconditional CV's can be obtained from the same set of values θ_r. When income is additive and linear or logarithmic, there exists an explicit formula for the $\theta_r' s$ that will be exploited below.

4.3 The standard logit and the logsum formula

The logsum formula has been used extensively in transportation, location theory and more recently in industrial organization (IO) as a simple welfare measure for the logit model (see the recent survey of de long et al. 2007). We provide below an extension of the standard model for conditional choices (see Eq. (26)).

If $v(y) = (1/\mu)\, y$, with $\mu > 0$, we have $\delta_y(c) = (1/\mu)\, c$ and $\psi_{kk} = \mu\delta_k$. We get the following explicit expression of the integral term

$$\theta_r = \mu\left(\frac{\delta_{r+1} - \delta_r + \ln\Omega_{r+1} - \ln\Omega_r}{s_r}\right),\ r = 1...n-1.$$

Using these expression of θ_r in (30) leads to the following formula for the unconditional expected CV:

$$\mathbb{E}[cv] = \mu\ln(\sigma_0/s_n) = \mu\ln\left(\sum_k e^{\omega_k} / \sum_k e^{v_k}\right). \tag{31}$$

This expression (31) corresponds to the difference between the ex-post and ex-ante logsums. The well known log-sum formula has been derived by McFadden as a welfare measure. The formula for the conditional CV's (see Proposition 10 and Corollary 11) are explicit in this case. Our analysis allows to compute conditional logsums which provide more accurate evaluation of surplus when ex-ante and/or ex-post choices are observed (see the numerical evaluations provided in Section 4.1). The reader is also referred to de Palma and Kilani (2007) who focus on characterizations of the conditional (to the ex-ante) distributions of maximum utility in RUM.

4.4 An alternative welfare measure for the CES

When the utility is additive but non linear in income, as for the CES model, we can still derive an explicit formula for the expected CV's. If $v(y) = (1/\mu)\ln y$, with $\mu > 0$, we have $\delta_y(c) = -(1/\mu)\ln(1 - c/y)$ and $\psi_{kk} = y\left(1 - e^{-\mu\delta_k}\right)$. The integral term in this case is given by (Use the change of variable $t = s_r / \left[s_r + \sigma_r (1 - c/y)^{1/\mu}\right]$)

$$\theta_r = \mu y \frac{s_r^{\mu-1}}{\sigma_r^{\mu}} B_{\frac{s_r}{\Omega_r}, \frac{s_r}{\Omega_{r+1}}}(1-\mu, \mu),\ r = 1...n-1,$$

where B denotes the generalized incomplete Beta function. The expected CV for the logarithmic in income logit model is

$$\mathbb{E}[cv] = y\left[1 - e^{-\mu\delta_n} - \frac{1}{\beta}\sum_{r=1}^{n-1}\left(\frac{s_r}{\sigma_r}\right)^{\mu} B_{\frac{s_r}{\Omega_r},\frac{s_r}{\Omega_{r+1}}}(1-\mu,\mu)\right]. \tag{32}$$

Assume for example that the systematic component of the utility has the following specification: $v_k = (1/\mu)(\ln y - \ln p_k)$ where p_k denotes the ex-ante price of good k. Using the Roy's identity, the ex-ante demand for good i is:

$$\mathbb{X}_i = \frac{p_i^{-\frac{1}{\mu}-1}}{\sum_k p_k^{-\frac{1}{\mu}}} y.$$

Anderson, de Palma and Thisse (1987) have shown that the CES representative consumer model (see Dixit and Stiglitz 1977) can be derived as a logit model with income additive logarithmic specification and double-exponentially distributed error terms. We provide below an expression for the conditional (and unconditional) CV corresponding to the CES. Anderson et al. (1992, pp. 97-100) show that "a rise in the CES indirect utility function does not necessarily imply that all constituent consumers (...) can be made better off by appropriate redistribution of income." This criticism of the representative consumer can be handled when the CV is computed at the individual level and *then* aggregated over the population. We provide this result below. Consider a change in prices $(p_1 ... p_n) \rightarrow (\rho_1 ... \rho_n)$, where ρ_k is the ex-post price of good k. In this case, the expected (aggregated) CV for the CES is given by

$$\mathbb{E}[cv] = y\left[1 - \frac{\rho_n}{p_n} - \mu\sum_{r=1}^{n-1}\frac{\Pi_r}{P_r} \times B_{\frac{s_r}{\Omega_r},\frac{s_r}{\Omega_{r+1}}}(1-\mu,\mu)\right], \tag{33}$$

where $P_r = \left(\sum_{k=1}^{r} p_k^{-1/\mu}\right)^{-\mu}$ and $\Pi_r = \left(\sum_{k=r+1}^{n} \rho_k^{-1/\mu}\right)^{-\mu}$ are respectively the partial ex-ante and the ex-post CES price indices, and where in this case the arguments of the Beta function are such that:

$$\begin{cases} s_r/\Omega_r = \left[1 + (p_r\Pi_r/\rho_r P_r)^{-1/\mu}\right]^{-1}, \\ s_r/\Omega_{r+1} = \left[1 + \left(p_{r+1}\Pi_r/\rho_{r+1}P_r\right)^{-1/\mu}\right]^{-1}. \end{cases} \tag{34}$$

These expressions differ from the aggregate standard welfare measures oh the CES model. They provide alternative welfare measure to assess the policy implication of price changes.

5 Concluding remarks

In this paper, we have presented a first step towards a dynamic choice model, where individuals may alter their current choice after a change in the attributes of the alternatives. For ARUM, we have computed the transition choice probabilities and the associated welfare measures (CV) and have provided analytical functional forms. Using these formulae will ease the welfare analysis both at the theoretical and empirical levels.

The proposed framework can be extended in several dimensions. The most important extension involves the mixed logit model widely used in empirical applications (see Berry, Levinsohn and Pakes 1999). In this case, some parameters are distributed so that the transition choice probabilities will involve a kernel that we have computed in Section 2, while the various welfare measures (conditional and unconditional distribution and moments of CV) will involve kernels provided in Section 4. In this sense, the mixed logit would only add an integral for each of the parameters that are being distributed.

We have concentrated our analysis on the case where only one series of change occur at once, and individual choices are observed ex-ante and ex-post (i.e. before and after this change). Moreover, we have assumed that the error terms remain the same, but this is not necessary the case in a truly dynamic model. It is easy to consider situations, and model situations where individuals have some probability to inherit a new error term (for some alternatives) when a change has occurred. Besides, practical situations may involve several changes staggered over time. In this case, the exact dynamics of the error term is relevant. Indeed, without change of the error terms, each change induces transitions which provide information on the parameters of the systematic utility as well as on the value of the error terms. As a consequence the model may lead to inconsistent sequence of choice if the error terms are individual specific. The redraw of the error terms allows to avoid these inconsistent situations. There is still a long way to compute exact formulae for truly dynamic random utility models. We hope that this paper provides a useful first step in these directions.

Acknowledgments: We would like to thank Simon Anderson, Moshe Ben-Akiva, Denis Bolduc, Raymond Deneckere, Nicolas Gravel, Hervé Moulin, Daniel McFadden, Robin Lindsey, Eric Renault, Jacques Thisse and Kenneth Train for comments and suggestions on various stages of this paper. The first author would like to thank the Institut Universitaire de France.

Appendix. Proofs

Proof of Theorem 1. The probability $\mathbb{P}_{i\hookrightarrow i}$ (see Eq. (3)) given by $\mathbb{P}_{i\hookrightarrow i} = \Pr\left(U_i > U_k, k \neq i; \Upsilon_i > \Upsilon_r, r \neq i\right)$, can be rewritten as

$$\mathbb{P}_{i\hookrightarrow i} = \Pr\left(U_i > U_k, k \neq i; U_i > U_r + \left(\delta_r - \delta_i\right), r \neq i\right),$$

and further simplified as

$$\mathbb{P}_{i\hookrightarrow i} = \Pr\left(U_i > U_k + \left(\delta_k - \delta_i\right)^+, k \neq i\right). \tag{35}$$

Comparing (35) with (1), we deduce that

$$\mathbb{P}_{i\hookrightarrow i} = \mathbb{P}_i\left(v_1 + \left(\delta_1 - \delta_i\right)^+ ... v_n + \left(\delta_n - \delta_i\right)^+\right).$$

If $j \neq i$, with $\delta_j > \delta_i$, $\mathbb{P}_{i\hookrightarrow j}$ given by (3) can be rewritten as

$$\mathbb{P}_{i\hookrightarrow j} = \Pr\left(U_i > U_k + \left(\delta_k - \zeta_{ij}\right)^+, k \neq i, j; \delta_j > \zeta_{ij} > \delta_i\right), \tag{36}$$

where the random variable $\zeta_{ij} \equiv \Upsilon_j - U_i$ represents the utility variation ex-post. Clearly, if $i > j$ and therefore $\delta_i \geq \delta_j$, then $\mathbb{P}_{i\hookrightarrow j} = 0$ as required.
If $j > i$, we associate to U_i and to Υ_j the variables of integration u and w, respectively. Remark that if $z \equiv w - u$ verifies $\delta_j \geq z \geq \delta_i$, then $u - v_i = u - v_i - (\delta_i - z)^+$ and $w - \omega_j = u - v_j - (\delta_j - z)^+$. The transition choice probability (36) can then be written in the following integral form:

$$\mathbb{P}_{i\hookrightarrow j} = \int_{-\infty}^{\infty}\int_{u+\delta_i}^{u+\delta_j} F^{ij}\left(u - v_1 - (\delta_1 - z)^+ ... u - v_n - (\delta_n - z)^+\right) dwdu.$$

Using the change of variable $z = w - u$ within the inner integral, we get:

$$\mathbb{P}_{i\hookrightarrow j} = \int_{-\infty}^{\infty}\int_{\delta_i}^{\delta_j} F^{ij}\left(u - v_1 - (\delta_1 - z)^+ ... u - v_n - (\delta_n - z)^+\right) dzdu.$$

The Fubini's theorem allows us to permute the integral signs so that:

$$\mathbb{P}_{i\hookrightarrow j} = \int_{\delta_i}^{\delta_j}\int_{-\infty}^{\infty} F^{ij}\left(u - v_1 - (\delta_1 - z)^+ ... u - v_n - (\delta_n - z)^+\right) dudz.$$

Thanks to Eq. (2), the inner integral is $\Pi_i^j\left(v_1 + (\delta_1 - z)^+ ... v_n + (\delta_n - z)^+\right)$, and therefore:

$$\mathbb{P}_{i\hookrightarrow j} = \int_{\delta_i}^{\delta_j} \Pi_i^j\left(v_1 + (\delta_1 - z)^+ ... v_n + (\delta_n - z)^+\right) dz,$$

which is the required expression. □

Proof of Proposition 2. If $j = i$, using Eq. (5) with the logit choice probabilities (8) we get $\mathbb{P}_{i\hookrightarrow i} = e^{v_i}/\Omega_i$, where $\Omega_i = \sum_{k\leq i} e^{v_k} + \sum_{k>i} e^{\omega_k - \delta_i}$, $i < n$

and where $s_n = \sum_k e^{v_k}$.
If $\delta_j > \delta_i$, with $n > j > i > 1$, using Eq. (4), we have:

$$\mathbb{P}_{i \hookrightarrow j} = \sum_{r=i}^{j-1} \int_{\delta_r}^{\delta_{r+1}} \Pi_i^j (v_1 ... v_r, \omega_{r+1} - z ... \omega_n - z) dz.$$

For the logit, $\Pi_i^j = \mathbb{P}_i \mathbb{P}_j$ so that

$$\mathbb{P}_{i \hookrightarrow j} = e^{v_i} e^{\omega_j} \sum_{r=i}^{j-1} \int_{\delta_r}^{\delta_{r+1}} \frac{e^{-z}}{\left(s_r + \sigma_r e^{-z}\right)^2} dz.$$

We integrate in each interval $[\delta_r, \delta_{r+1}]$ to get:

$$\mathbb{P}_{i \hookrightarrow j} = \sum_{r=i}^{j-1} \left(\frac{e^{v_i}}{\Omega_{r+1}} - \frac{e^{v_i}}{\Omega_r} \right) \frac{e^{\omega_j}}{\sigma_r},$$

since $s_r + \sigma_r e^{-\delta_r} = \Omega_r$ and

$$\begin{array}{c} s_r + \sigma_r e^{-\delta_{r+1}} = \sum_{k \leq r} e^{v_k} + \sum_{k > r} e^{\omega_k - \delta_{r+1}} = \\ (s_r - e^{v_{r+1}}) + \left(e^{\omega_{r+1} - \delta_{r+1}} + \sigma_r e^{-\delta_{r+1}}\right) = \Omega_{r+1}. \end{array}$$

The remaining cases $i = 1$ and $j = n$ are left to the reader. $\square$
Proof of Lemma 3. First note that ψ_{ii} restores the utility of i to its ex-ante level U_i, since $\Upsilon_i (y - \psi_{ii}) = \delta_i (\psi_{ii}) + U_i = U_i$.
For a transition $i \hookrightarrow i$, we have $U_i \geq U_k + (\delta_k - \delta_i)^+$ (see (35)). As a consequence, since $\Upsilon_k (y - c) = U_k + \delta_k (c)$, then ψ_{ik} (which solves $\delta_k (\psi_{ik}) = (\delta_k - \delta_i)^+$) is the largest amount needed to restore the utility of alternative k to the ex-ante level U_i. As a consequence, $\psi_{ii} = m_i \leq cv \leq \max_k (\psi_{ik}) = M_i$.
For a transition $i \hookrightarrow j$, $j > i$, since $U_j + (\delta_j - \delta_i) \geq U_i \geq U_j$, then ψ_{ij} (which solves $\delta_j \left(\psi_{ij}\right) = \delta_j - \delta_i$) and ψ_{jj} (which solves $\delta_j \left(\psi_{jj}\right) = 0$) are respectively the lowest and the largest amount needed to restore the utility of alternative j to the ex-ante level U_i, with necessarily $\psi_{ij} \leq \psi_{jj}$. Moreover, for $k \neq i$, j, we have $U_k + \left(\delta_k - \xi_{ij}\right)^+ \leq U_i$, where $\xi_{ij} \equiv \Upsilon_j - U_i$ (see (36)). Since $\delta_j \geq \xi_{ij} \geq \delta_i$, ψ_{jk} (which solves $\delta_j \left(\psi_{jj}\right) = (\delta_k - \delta_j)^+$) is the largest amount needed to restore the utility of alternative k to the ex-ante level U_i. Altogether, the above conditions imply:

$$\max \left(\psi_{ii}, \psi_{ij}\right) = m_{ij} \leq cv \leq \max \left[\psi_{ii}, \max_{k \neq i} \left(\psi_{jk}\right)\right].$$

Since $\delta_i \leq \delta_j$, we have that $\psi_{ji} = \psi_{ii}$, we get: $m_{ij} \leq cv \leq \overline{m}_j$. $\square$
Proof of Theorem 4. If i is chosen ex-ante, the event $\{c \geq cv\}$ can also be written as:

$$\left\{ \max_k \left[\Upsilon_k (y - cv)\right] \geq \max_k \left[\Upsilon_k (y - c)\right] \right\} = \left\{U_i \geq \Upsilon_k (y - c), \forall k\right\},$$

using the fact that the Υ_k's are strictly increasing in y and recalling the definition of cv. For $c \geq m_{ij} \geq \psi_{ii}$, we have necessarily:
$\Upsilon_i(y - \psi_{ii}) = U_i \geq \Upsilon_i(y - c)$, so we get $\{c \geq cv\} = \{U_i \geq \Upsilon_k(y - c), k \neq i\}$ or $\{c \geq cv\} = \{U_i \geq U_k + \delta_k(c), k \neq i\}$. Hence:

$$\{c \geq cv\} = \{U_i > U_k + \delta_k(c), k \neq i\}, \text{ a.e.,}$$

and we rewrite Eq. (13) as:

$$\Phi_{i \hookrightarrow j}(c) = \frac{\Pr(U_i > U_k + \delta_k(c), k \neq i; U_i > U_k, k \neq i; \Upsilon_j > \Upsilon_r, r \neq j)}{\mathbb{P}_{i \hookrightarrow j}(\underline{v}, \underline{\omega})},$$

or further as:

$$\Phi_{i \hookrightarrow j}(c) = \frac{\Pr\left(U_i > U_k + \delta_k^+(c), k \neq i; \Upsilon_j > \Upsilon_r, r \neq j\right)}{\mathbb{P}_{i \hookrightarrow j}(\underline{v}, \underline{\omega})}. \tag{37}$$

Comparing the numerator of Eq. (37) with Eq. (3), we deduce that it takes the form of a transition probability of the type $i \hookrightarrow j$ corresponding to a change $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right) \rightarrow \underline{\omega}$. Therefore, according to Theorem 4, we get Eq. (14).
According to Lemma 3, the support of cv conditional to transition $i \hookrightarrow j$ is included in $[m_{ij}, \overline{m}_j]$. We proof here that the support is $(m_{ij}, \overline{m}_j]$.
First, the ith component of $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right)$ is v_i while the other components are $v_k + \delta_k^+(c) \geq v_k$, $k \neq i$, with at least one strict inequality. As a consequence,

$$\mathbb{P}_{i \hookrightarrow j}(\underline{v}, \underline{\omega}) > \mathbb{P}_{i \hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right),$$

so that $1 > \Phi_{i \hookrightarrow j}(c)$. Therefore, the support of cv extends up to $\overline{m}_j$.
Second, if $j = i$, and $c \geq m_{ii} = \psi_{ii}$, we necessarily have $\Phi_{i \hookrightarrow j}(c) > 0$ since we always have

$$\mathbb{P}_{i \hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right) > 0.$$

Third, if $j > i$ (and $\delta_j > \delta_i$), let $c > m_{ij}$. We have $\delta_i^+(c) = 0$ since $c > \psi_{ii}$ and $\delta_j^+(c) < \delta_j - \delta_i$ since $c > \psi_{ij}$. As a consequence, in both cases, a transition $i \hookrightarrow j$ is feasible with a change $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right) \rightarrow \underline{\omega}$ (see Theorem 1) since

$$\omega_i - \left(v_i + \delta_i^+(c)\right) = \delta_i < \omega_j - \left(v_j + \delta_j^+(c)\right) = \delta_j - \delta_j^+(c), \tag{38}$$

which implies that $\Phi_{i \hookrightarrow j}(c) > 0$. Finally, note that if $m_{ij} = \psi_{ij}$, the previous inequality (38) became an equality for $c = \psi_{ij}$ so that the (conditional on $i \hookrightarrow j$) distribution of cv has no jump at the lower bound of the support, i.e. for $c = m_{ij}$. Otherwise, if $m_{ij} = \psi_{ii} > \psi_{ij}$, the inequality is still strict for $c = m_{ij} = \psi_{ii}$, so that the distribution has no jump at this point. $\square$
Proof of Corollary 5. (a) Using Theorem 4, for feasible transitions, we have $m_{ij} \geq \psi_{ii}$. Moreover, since ψ_{jk} solves $\delta_k\left(\psi_{jk}\right) = (\delta_k - \delta_j)^+$, and since $\delta_k(c)$ is decreasing in c, we have (recall that $\overline{m}_j \equiv \max_k\left(\psi_{jk}\right)$) the ranking:

$$\overline{m}_1 \leq ... \leq \overline{m}_n.$$

Since $\delta_k(\psi_{nk}) = (\delta_k - \delta_n)^+ = 0$, we have $\psi_{nk} = \psi_{kk}$ so that $\overline{m}_n = \max_k(\psi_{kk})$. Therefore, the support of cv conditional to the ex-ante choice of i is $[\psi_{ii}, \overline{m}_n]$. Moreover, according to Theorem 4, we get that:

$$\Phi_{i\hookrightarrow}(c) = \frac{\sum_{j\in\mathcal{F}_{i\hookrightarrow}} \mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right)}{\mathbb{P}_i(\underline{v})}, \tag{39}$$

where $\mathcal{F}_{i\hookrightarrow}$ stands for the set of alternatives j such that $i \hookrightarrow j$ is feasible. For non-feasible transitions $i \hookrightarrow j$ where $\delta_i \geq \delta_j$, if $c \geq \psi_{ii}$ the ith component of $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right)$ is v_i while its jth component is $v_j + \delta_j^+(c)$. We have

$$\omega_i - v_i = \delta_i \geq \omega_j - \left(v_j + \delta_j^+(c)\right) = \delta_j - \delta_j^+(c),$$

so that for a change $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right) \to \underline{\omega}$, the transitions $i \hookrightarrow j$ is non-feasible. Therefore,

$$\mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right) = 0.$$

This allows us to extent the sum sign in (39) to all alternatives to get:

$$\Phi_{i\hookrightarrow}(c) = \frac{\sum_j \mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right)}{\mathbb{P}_i(\underline{v})}.$$

Then, using Eq. (6), we get Eq. (15).
(b) According to Theorem 4, the support of cv conditional to the ex-post choice of j is $\left[\min_{i\in\mathcal{F}_j}(m_{ij}), \overline{m}_j\right]$ where $\mathcal{F}_j$ is the set of alternatives i such that $i \hookrightarrow j$ is feasible. For non feasible transitions verifying $\delta_i \geq \delta_j$, we have $\psi_{ij} = \psi_{jj}$ and therefore that $m_{ij} \geq \psi_{jj} = m_{jj}$. As a consequence, $\min_{i\in\mathcal{F}_j}(m_{ij}) = \min_i(m_{ij}) = \underline{m}_j$ and the support is $\left[\underline{m}_j, \overline{m}_j\right]$.
For $c \geq \underline{m}_j$, using Theorem 4, we get that

$$\Phi_{\hookrightarrow j}(c) = \frac{\sum_{i\in\mathcal{F}_j} H_{m_{ij}}(c) \times \mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right)}{\mathbb{P}_j(\underline{\omega})}. \tag{40}$$

The sum can be extended to non feasible transitions $i \hookrightarrow j$ to get Eq. (16). Indeed, either $c < m_{ij}$ and therefore $H_{m_{ij}}(c) = 0$ or, if $c \geq m_{ij}$, since $\delta_j\left(\psi_{ij}\right) = (\delta_j - \delta_i)^+ = 0$, we have that $c \geq m_{ij} \geq \psi_{jj}$. The ith component of $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right)$ is v_i and its jth component is v_j so that for $\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c)\right) \to \underline{\omega}$, the transitions $i \hookrightarrow j$ is non-feasible and hence $\mathbb{P}_{i\hookrightarrow j}\left(v_1 + \delta_1^+(c) ... v_n + \delta_n^+(c); \underline{\omega}\right) = 0$. $\square$

Proof of Theorem 6. For $0 \leq \pi \leq 1$, define the conditional quantile function $\mathbf{\Phi}_{i\hookrightarrow j}^{-1}(\pi) \equiv \sup\left\{c \in [m_{ij}, \overline{m}_j] \,|\, \pi \geq \mathbf{\Phi}_{i\hookrightarrow j}(c)\right\}$, which is the inverse of the conditional CDF of cv. By definition, the pth conditional moment of cv verifies $\mathbb{E}_{i\hookrightarrow j}[cv^p] \equiv \int_0^1 \left[\mathbf{\Phi}_{i\hookrightarrow j}^{-1}(\pi)\right]^p d\pi$. For $c \in [m_{ij}, \overline{m}_j]$, the functions $\mathbf{\Phi}_{i\hookrightarrow j}(c)$ is continuous and monotonic. It is therefore a.e. differentiable according to the Lebesgue theorem (cf. Rudin 1986). As a consequence, a PDF $\boldsymbol{\phi}_{i\hookrightarrow j}$ can a.e.

be defined. Using the change of variable: $\pi = \mathbf{\Phi}_{i\hookrightarrow j}(c)$, with:
$c \in [m_{ij}, \overline{m}_j]$, we get $\mathbb{E}_{i\hookrightarrow j}[cv^p] = m_{ij}^p \Phi_{i\hookrightarrow j}(m_{ij}) + \int_{m_{ij}}^{\overline{m}_j} z^p \boldsymbol{\phi}_{i\hookrightarrow j}(c)\,dc$. Then using an integration by parts, we obtain:

$$\mathbb{E}_{i\hookrightarrow j}[cv^p] = \overline{m}_j^p - p\int_{m_{ij}}^{\overline{m}_j} c^{p-1}\mathbf{\Phi}_{i\hookrightarrow j}(c)\,dc,$$

which coincides with (18). $\square$

Proof of Corollary 7. This proof uses the same technique as for the proof of Theorem 6 by considering $\Phi_{i\hookrightarrow}$ given by (15) instead of $\Phi_{i\hookrightarrow j}$ or by considering $\Phi_{\hookrightarrow j}$ given by (16) instead of $\Phi_{i\hookrightarrow j}$. $\square$

Proof of Proposition 8. Recall that (see Eq. (18)):

$$\mathbb{E}_{1\hookrightarrow j}[cv] = \overline{m}_j - \frac{1}{\mathbb{P}_{1\hookrightarrow j}(\underline{v}; \omega_1, v_2...v_n)}\int_{m_{1j}}^{\overline{m}_j} I_j(\delta_1, c)\,dc, \tag{41}$$

where $I_j(\delta_1, c) \equiv \mathbb{P}_{1\hookrightarrow j}\left(v_1 + \delta_1^+(c)...v_n + \delta_n^+(c); \omega_1, v_2...v_n\right)$, $j = 1...n$, and where $\overline{m}_j = \max_k\left(\psi_{jk}\right)$, with ψ_{jk} solving $\delta_k\left(\psi_{jk}\right) = (\delta_k - \delta_j)^+$, $k = 1...n$, (see Eq. (12)).

Note that $\psi_{11} < 0$ since $\delta_1 < 0$. The Roy's Identity applied in the deterministic case leads to: $\lim_{\Delta p_1 \to 0^+}(\psi_{11}/\Delta p_1) = -x_1$. Moreover, since $\delta_k = 0$, $k = 2...n$, we have: $\delta_k(\psi_{1k}) = (0 - \delta_1)^+ = -\delta_1$, $k = 2...n$. Accordingly, $\psi_{1k} < 0$, $k = 2...n$, and $\lim_{\delta_1 \to 0^-}(\psi_{1k}/\delta_1) = (\partial V_k/\partial y)^{-1}$. Therefore, using again the Roy's Identity in the deterministic case we have:

$$\lim_{\Delta p_1 \to 0^+}\left(\frac{\psi_{1k}}{\Delta p_1}\right) = -x_1\rho_{1k}, k = 1...n.$$

Therefore: $\lim_{\Delta p_1 \to 0^+}(\overline{m}_1/\Delta p_1) = -\min_k(\rho_{1k})\,x_1$. Now, since $I_1(\delta_1, c)$ is continuous in c, using the mean value theorem for integration, we get

$$\mathbb{E}_{1\hookrightarrow 1}[cv] = \overline{m}_1 - \frac{(\overline{m}_1 - \psi_{11})\,I_1(\delta_1, \widetilde{c}_1)}{\mathbb{P}_{1\hookrightarrow 1}(\underline{v}; \omega_1, v_2...v_n)},$$

where $\widetilde{c}_1 \in (\psi_{11}, \overline{m}_1)$. Now using the fact that $\lim_{\Delta p_1 \to 0^+} I_1(\delta_1, \widetilde{c}_1) = I_1(0, 0)$ $= \mathbb{P}_{1\hookrightarrow 1}(\underline{v}, \underline{v}) = \mathbb{P}_1(\underline{v})$ and that $\lim_{\Delta p_1 \to 0^+}\mathbb{P}_{1\hookrightarrow 1}(\underline{v}; \omega_1, v_2...v_n) = \mathbb{P}_1(\underline{v})$, we get:

$$\lim_{\Delta p_1 \to 0^+}\frac{\mathbb{E}_{1\hookrightarrow 1}[cv]}{\Delta p_1} = -\min_k(\rho_{1k})\,x_1 - \left(-\min_k(\rho_{1k})\,x_1 - x_1\right) = -x_1.$$

Let $j > 1$. Since $\delta_k \leq 0$, $k = 1...n$, and $\delta_j = 0$, we have: $\delta_k\left(\psi_{jk}\right) = \delta_k^+ = 0$. As a consequence, $\psi_{j1} = \psi_{11} < 0$ (since $\delta_1(\psi_{11}) = \delta_1\left(\psi_{j1}\right) = 0$) and $\psi_{jk} = 0$, $k > 1$. Hence, $\overline{m}_j = 0$ which allow us to rewrite (41) as:

$$\mathbb{E}_{1\hookrightarrow j}[cv] = \frac{1}{\mathbb{P}_{1\hookrightarrow j}(\underline{v}; \omega_1, v_2...v_n)}\int_0^{m_{1j}} I_j(\delta_1, c)\,dc.$$

Using Eq. (4) and applying the mean value theorem for integration we get:

$$\mathbb{E}_{1\hookrightarrow j}\left[cv\right] = -\frac{\int_0^{m_{1j}} I_j\left(\delta_1, c\right) dc}{\delta_1 \Pi_1^j\left(v_1, v_2 - \widetilde{\delta} ... v_n - \widetilde{\delta}\right)},$$

where $\widetilde{\delta} \in (\delta_1, 0)$. Using Eq. (4) we rewrite $I_j\left(\delta_1, c\right)$ as:

$$I_j\left(\delta_1, c\right) = \int_{\delta_1}^{-\delta_j(c)} \Pi_1^j\left(v_1, v_2 + \left(-\delta_2\left(c\right) - z\right)^+ ... v_n + \left(-\delta_n\left(c\right) - z\right)^+\right) dz.$$

Let $\varepsilon > 0$ small enough. Since the integrand tends towards Π_1^j as δ_1 and z tend towards zero, we can find δ_1 and c arbitrarily small in order that

$$\left(-\delta_j\left(c\right) - \delta_1\right)\left(\Pi_1^j - \varepsilon\right) \leq I_j\left(\delta_1, c\right) \leq \left(-\delta_j\left(c\right) - \delta_1\right)\left(\Pi_1^j + \varepsilon\right).$$

Applying the Taylor's theorem to $\delta_j\left(c\right)$, we get

$$\left(\frac{\partial V_j}{\partial y} c - R - \delta_1\right)\left(\Pi_1^j - \varepsilon\right) \leq I_j\left(\delta_1, c\right) \leq \left(\frac{\partial V_j}{\partial y} c - R - \delta_1\right)\left(\Pi_1^j + \varepsilon\right),$$

where R verifies $|R| \leq Mc^2$ with M a positive constant. Therefore, by integration and taking the limit $\varepsilon \to 0$, we get:

$$\lim_{\delta_1 \to 0^-} \frac{1}{\delta_1^2} \int_0^{m_{1j}} I_j\left(\delta_1, c\right) dc = -\left(l_j - \frac{\partial V_j}{\partial y}\frac{l_j^2}{2}\right)\Pi_1^j.$$

where $l_j \equiv \lim_{\delta_1 \to 0^-}\left(m_{1j}/\delta_1\right)$. Recall that $m_{1j} = \max\left(\psi_{11}, \psi_{1j}\right)$. Now, using the chain rule, we get:

$$l_j = \min\left(\lim_{\delta_1 \to 0^-} \frac{\psi_{11}}{\delta_1}, \lim_{\delta_1 \to 0^-} \frac{\psi_{1j}}{\delta_1}\right) = \left\{\begin{array}{ll} \left(\partial V_j/\partial y\right)^{-1}, & \text{if } \rho_{1j} \leq 1; \\ \left(\partial V_1/\partial y\right)^{-1}, & \text{if } \rho_{j1} \leq 1. \end{array}\right. \quad (42)$$

Using again the chain rule and the Roy's Identity, we get:

$$\lim_{\Delta p_1 \to 0^+} \frac{\mathbb{E}_{1\hookrightarrow j}\left[cv\right]}{\Delta p_1} = \frac{x_1\left(\partial V_1/\partial y\right)}{\Pi_1^j} \lim_{\delta_1 \to 0^-} \frac{1}{\delta_1^2} \int_0^{m_{1j}} I_j\left(\delta_1, c\right) dc.$$

Hence

$$\lim_{\Delta p_1 \to 0^+} \frac{\mathbb{E}_{1\hookrightarrow j}\left[cv\right]}{\Delta p_1} = \left\{\begin{array}{ll} -\frac{\rho_{1j}}{2} x_1, & \text{if } \rho_{1j} \leq 1; \\ -\left(1 - \frac{\rho_{j1}}{2}\right) x_1, & \text{if } \rho_{j1} \leq 1. \end{array}\right.$$

Now, recall that (see Eq. (18)):

$$\mathbb{E}_{i\hookrightarrow n}\left[cv\right] = \overline{m}_n - \frac{1}{\mathbb{P}_{i\hookrightarrow n}\left(\underline{v}; \omega_1, v_2 ... v_n\right)} \int_{m_{in}}^{\overline{m}_n} J_i\left(\delta_n, c\right) dc, \quad (43)$$

where $J_i\left(\delta_n, c\right) \equiv \mathbb{P}_{i\hookrightarrow n}\left(v_1 + \delta_1^+\left(c\right) ... v_n + \delta_n^+\left(c\right); v_1 ... v_{n-1}, \omega_n\right)$, $j = 1...n$, and where $\overline{m}_n = \max_k\left(\psi_{nk}\right)$, with ψ_{nk} solving $\delta_k\left(\psi_{nk}\right) = \left(\delta_k - \delta_n\right)^+$, with $k =$

$1...n$, (see Eq. (12)). Since $\delta_k \leq \delta_n$, then ψ_{nk} is solving $\delta_k(\psi_{nk}) = 0$, with $k = 1...n$, (see Eq. (12)). Therefore, $\psi_{nk} = 0$, $k = 1...n-1$ and $\psi_{nn} > 0$. Therefore, $\overline{m}_n = \max_k(\psi_{nk}) = \psi_{nn}$. Moreover, we have:

$$m_{in} = \max(\psi_{ii}, \psi_{in}) = \max(0, \psi_{in}), \; i = 1...n-1,$$

where ψ_{in} is solving $\delta_n(\psi_{in}) = (\delta_n - \delta_i)^+ = \delta_n$. Therefore, $\psi_{in} = 0$ and $m_{in} = 0$. For $c \in (0, \psi_{nn})$, we have

$$\begin{aligned} J_i(\delta_n, c) &= \int_0^{\delta_n - \delta_n(c)} \Pi_i^n(v_1..., v_{n-1}, \omega_n - z)\, dz \\ &= (\delta_n - \delta_n(c))\, \Pi_i^n(v_1..., v_{n-1}, \omega_n - \widetilde{z}), \end{aligned}$$

where $\widetilde{z} \in (0, \delta_n - \delta_n(c))$. Using the fact that $\Pi_i^n(v_1..., v_{n-1}, \omega_n - \widetilde{z})$ tends towards Π_i^n as δ_n tends towards zero and applying the Taylor's theorem to $\delta_n(c)$, we get:

$$\lim_{\delta_n \to 0^+} \frac{1}{\delta_n^2} \int_0^{\psi_{nn}} J_i(\delta_1, c)\, dc = \frac{\Pi_i^n}{2} \frac{\partial V_n}{\partial y} \lim_{\delta_n \to 0^+} \frac{\psi_{nn}^2}{\delta_n^2} = \frac{\Pi_i^n}{2} (\partial V_n / \partial y)^{-1}.$$

Therefore, using the chain rule, we get:

$$\lim_{\Delta p_n \to 0^-} \frac{\mathbb{E}_{i \hookrightarrow n}[cv]}{\Delta p_n} = \frac{(\partial V_n / \partial p_n)}{2(\partial V_n / \partial y)} = -\frac{x_n}{2}$$

Now, since $\mathbb{E}_{n \hookrightarrow n}[cv] = \psi_{nn}$, we have $\lim_{\Delta p_n \to 0^-}(\mathbb{E}_{n \hookrightarrow n}[cv] / \Delta p_n) = -x_n$. $\square$

Proof of Proposition 9. We have: $\delta_k(c) = \delta_k - \delta_y(c)$, where we define: $\delta_y(c) \equiv v(y) - v(y - c)$, is strictly increasing in c. The $\psi'_{ik}s$, defined by (12), verify:

$$\psi_{ik} = \begin{cases} \delta_y^{-1}(\delta_k), & \text{if } k < i; \\ \delta_y^{-1}(\delta_i), & \text{if } k \geq i. \end{cases}$$

Note that $\psi_{ik} \leq \psi_{ii}$ since δ_y^{-1} is increasing and since $\delta_k \leq \delta_i$ for $k \leq i$. therefore, the support of the distribution of cv conditional on the transition $i \hookrightarrow j$, $j \geq i$, is: $[\psi_{ii}, \psi_{jj}]$, since $\psi_{ij} = \psi_{ii} = \delta_y^{-1}(\delta_i) = m_{ij}$ and since $\overline{m}_j = \delta_y^{-1}(\delta_j) = \psi_{jj}$. If $c \in \left[\psi_{ll}, \psi_{(l+1)(l+1)}\right]$, then $\underline{v} + \underline{\delta}^+(c) = (v_1...v_l, \omega_{l+1} - \delta_y(c)...\omega_n - \delta_y(c))$ so that $\underline{\omega} - [\underline{v} + \underline{\delta}^+(c)] = (\delta_1...\delta_i...\delta_l, \delta_y(c)...\delta_y(c))$. Therefore, we have the ranking $\delta_1 \leq ... \leq \delta_l \leq \delta_y(c)$. Using Eq. (10) (see Proposition 2) we get $\mathbb{P}_{i \hookrightarrow j}(\underline{v} + \underline{\delta}^+(c)) = e^{v_i + \omega_j} \sum_{r=i}^{j-1} \sigma_r^{-1} \left[\Omega_{(r+1)l}^{-1}(c) - \Omega_{rl}^{-1}(c)\right]$, where

$$\Omega_{rl}(c) = \begin{cases} \Omega_r, & \text{if } r \leq l; \\ s_l + \sigma_l e^{-\delta_y(c)} & \text{if } r > l. \end{cases}$$

As a consequence, for $j > l \geq i$, we have

$$\mathbb{P}_{i \hookrightarrow j}(\underline{v} + \underline{\delta}^+(c)) = e^{v_i + \omega_j} \left\{ \Xi_{il} + \sigma_l^{-1} \left[\left(s_l + \sigma_l e^{-\delta_y(c)}\right)^{-1} - \Omega_l^{-1} \right] \right\}.$$

Using the fact that $\mathbb{P}_{i\hookrightarrow j}\left(\underline{v}\right) = e^{v_i} e^{\omega_j} \Xi_{ij}$, $j > i$, we get Eq. (25). $\square$

Proof of Proposition 10. Clearly, for a transition $i \hookrightarrow i$, we have $\mathbb{E}_{i\hookrightarrow i}\left[cv\right] = \psi_{ii}$. For a feasible transition $i \hookrightarrow j$, with $j > i$, using Theorem (6) with $p = 1$, we get

$$\mathbb{E}_{i\hookrightarrow j}\left[cv\right] = \psi_{jj} - \int_{\psi_{ii}}^{\psi_{jj}} \Phi_{i\hookrightarrow j}\left(c\right) dc,$$

which can be rewritten as

$$\mathbb{E}_{i\hookrightarrow j}\left[cv\right] = \psi_{jj} - \sum_{l=i}^{j-1} \int_{\psi_{ll}}^{\psi_{(l+1)(l+1)}} \Phi_{i\hookrightarrow j}\left(c\right) dc. \tag{44}$$

Using (25) and (44), we get

$$\mathbb{E}_{i\hookrightarrow j}\left[cv\right] = \psi_{jj} - \frac{1}{\Xi_{ij}} \sum_{l=i}^{j-1} \int_{\psi_{ll}}^{\psi_{(l+1)(l+1)}} \left[\Xi_{il} + \frac{1}{\sigma_l}\left(\frac{1}{s_l + \sigma_l e^{-\delta_y(c)}} - \frac{1}{\Omega_l}\right)\right] dc,$$

which can be rewritten as

$$\begin{aligned}
\mathbb{E}_{i\hookrightarrow j}\left[cv\right] &= \psi_{jj} - \frac{1}{\Xi_{ij}} \sum_{l=i}^{j-1} \int_{\psi_{ll}}^{\psi_{(l+1)(l+1)}} \Xi_{i(l+1)} dc \\
&\quad + \frac{1}{\Xi_{ij}} \sum_{l=i}^{j-1} \sigma_l^{-1} \left[\Omega_{l+1}^{-1}\left(\psi_{(l+1)(l+1)} - \psi_{ll}\right) - \theta_l\right].
\end{aligned}$$

Using the fact that $\Xi_{i(l+1)} = \sum_{r=i}^{l} \sigma_r^{-1}\left(\Omega_{r+1}^{-1} - \Omega_r^{-1}\right)$ and inverting the sign sums we get

$$\begin{aligned}
\mathbb{E}_{i\hookrightarrow j}\left[cv\right] &= \psi_{jj} - \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \int_{\psi_{rr}}^{\psi jj} \sigma_r^{-1}\left(\Omega_{r+1}^{-1} - \Omega_r^{-1}\right) dc \\
&\quad + \frac{1}{\Xi_{ij}} \sum_{l=i}^{j-1} \sigma_l^{-1} \left[\Omega_{l+1}^{-1}\left(\psi_{(l+1)(l+1)} - \psi_{ll}\right) - \theta_l\right],
\end{aligned}$$

which can be simplified

$$\begin{aligned}
\mathbb{E}_{i\hookrightarrow j}\left[cv\right] &= \psi_{jj} - \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \int_{\psi_{rr}}^{\psi jj} \sigma_r^{-1}\left(\Omega_{r+1}^{-1} - \Omega_r^{-1}\right) dc \\
&\quad + \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \sigma_r^{-1} \left[\Omega_{r+1}^{-1}\left(\psi_{(r+1)(r+1)} - \psi_{rr}\right) - \theta_r\right],
\end{aligned}$$

or further as

$$\begin{aligned}
\mathbb{E}_{i\hookrightarrow j}\left[cv\right] &= \psi_{jj} - \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \sigma_r^{-1}\left(\Omega_{r+1}^{-1} - \Omega_r^{-1}\right)\left(\psi_{jj} - \psi_{rr}\right) \\
&\quad + \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \int_{\psi_{rr}}^{\psi_{(r+1)(r+1)}} \sigma_r^{-1}\Omega_{r+1}^{-1} dc - \frac{1}{\Xi_{ij}} \sum_{lr=i}^{j-1} \sigma_r^{-1}\theta_r.
\end{aligned}$$

We further simplify this expression as:

$$\begin{aligned}
\mathbb{E}_{i\hookrightarrow j}\left[cv\right] &= \frac{1}{\Xi_{ij}}\sum_{r=i}^{j-1}\sigma_r^{-1}\left(\Omega_{r+1}^{-1}-\Omega_r^{-1}\right)\psi_{rr} \\
&+\frac{1}{\Xi_{ij}}\sum_{r=i}^{j-1}\sigma_r^{-1}\Omega_{r+1}^{-1}\left(\psi_{(r+1)(r+1)}-\psi_{rr}\right)-\frac{1}{\Xi_{ij}}\sum_{lr=i}^{j-1}\sigma_r^{-1}\theta_r,
\end{aligned}$$

which is equivalent to Eq. (26). $\square$

Proof of Corollary 11. (a) If $i<n$, we have

$$\mathbb{E}_{i\hookrightarrow}\left[cv\right]=\sum_{j=i}^{n}\frac{\mathbb{P}_{i\hookrightarrow j}\left(\underline{v},\underline{\omega}\right)}{\mathbb{P}_i\left(\underline{v}\right)}\mathbb{E}_{i\hookrightarrow j}\left[cv\right]. \tag{45}$$

Using (8) and (10) (see Proposition 2), the ratio of probabilities are such that

$$\frac{\mathbb{P}_{i\hookrightarrow j}\left(\underline{v},\underline{\omega}\right)}{\mathbb{P}_i\left(\underline{v}\right)}=\left\{\begin{array}{cl} s_n/\Omega_i, & \text{if } j=i; \\ s_n e^{\omega_j}\Xi_{ij}, & \text{if } j>i; \end{array}\right. \tag{46}$$

Therefore, using (46) and (26) (see Proposition 10), we write (45) as

$$\mathbb{E}_{i\hookrightarrow}\left[cv\right]=s_n\left\{\frac{\psi_{ii}}{\Omega_i}+\sum_{j=i+1}^{n}\sum_{r=i}^{j-1}\frac{e^{\omega_j}}{\sigma_r}\left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}}-\frac{\psi_{rr}}{\Omega_r}-\theta_r\right]\right\},$$

which can be rewritten by inverting the two sum signs as

$$\mathbb{E}_{i\hookrightarrow}\left[cv\right]=s_n\left\{\frac{\psi_{ii}}{\Omega_i}+\sum_{r=i}^{n-1}\sum_{j=r+1}^{n}\frac{e^{\omega_j}}{\sigma_r}\left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}}-\frac{\psi_{rr}}{\Omega_r}-\theta_r\right]\right\},$$

and simplified as

$$\mathbb{E}_{i\hookrightarrow}\left[cv\right]=s_n\left\{\frac{\psi_{ii}}{\Omega_i}+\sum_{r=i}^{n-1}\left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}}-\frac{\psi_{rr}}{\Omega_r}-\theta_r\right]\right\}.$$

This expression can be rewritten as in Eq. (28).
Finally, if $i=n$, clearly we have $\mathbb{E}_{n\hookrightarrow}\left[cv\right]=\psi_{nn}$.
(b) If $i=1$ clearly we have $\mathbb{E}_{\hookrightarrow 1}\left[cv\right]=\psi_{11}$.
If $j>1$, we have

$$\mathbb{E}_{\hookrightarrow j}\left[cv\right]=\sum_{i=1}^{j}\frac{\mathbb{P}_{i\hookrightarrow j}}{\mathbb{P}_j}\mathbb{E}_{i\hookrightarrow j}\left[cv\right]. \tag{47}$$

Using (8) and (10) (see Proposition 2), we get the ratio of probabilities

$$\frac{\mathbb{P}_{i\hookrightarrow j}}{\mathbb{P}_j}=\left\{\begin{array}{cl} \sigma_0 e^{-\delta_j}/\Omega_j, & \text{if } i=j; \\ \sigma_0 e^{v_i}\Xi_{ij}, & \text{if } i<\ j; \end{array}\right. \tag{48}$$

From (48) and (26) we get

$$\mathbb{E}_{\hookrightarrow j}\left[cv\right] = \sigma_0 \left\{ \sum_{i=1}^{j-1} \sum_{r=i}^{j-1} \frac{e^{v_i}}{\sigma_r} \left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}} - \frac{\psi_{rr}}{\Omega_r} - \theta_r \right] + \frac{\psi_{jj}}{e^{\delta_j}\Omega_j} \right\}.$$

Inverting the two sum signs we obtain

$$\mathbb{E}_{\hookrightarrow j}\left[cv\right] = \sigma_0 \left\{ \sum_{r=1}^{j-1} \sum_{i=1}^{r} \frac{e^{v_i}}{\sigma_r} \left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}} - \frac{\psi_{rr}}{\Omega_r} - \theta_r \right] + \frac{\psi_{jj}}{e^{\delta_j}\Omega_j} \right\},$$

which can be simplified as

$$\mathbb{E}_{\hookrightarrow j}\left[cv\right] = \sigma_0 \left\{ \sum_{r=1}^{j-1} \frac{s_r}{\sigma_r} \left[\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}} - \frac{\psi_{rr}}{\Omega_r} - \theta_r \right] + \frac{\psi_{jj}}{e^{\delta_j}\Omega_j} \right\}.$$

This expression can be rewritten as

$$\mathbb{E}_{\hookrightarrow j}\left[cv\right] = \sigma_0 \left\{ \sum_{r=2}^{j} \frac{s_r\psi_{rr}}{\sigma_r\Omega_r} - \sum_{r=1}^{j-1} \frac{s_r\psi_{rr}}{\sigma_r\Omega_r} - \sum_{r=1}^{j-1} \frac{s_r\theta_r}{\sigma_r} + \frac{\psi_{jj}}{\Omega_j e^{\delta_j}} \right\}.$$

Noting that $\sigma_r s_r - \sigma_r s_r = -e^{\omega_r}\Omega_r$, that $s_1/\sigma_1\Omega_1 = e^{\omega_1}/\sigma_0\sigma_1$ and moreover that: $s_{j-1} + e^{-\delta_j}\sigma_{j-1} = \Omega_j$, we obtain the required expression (29). $\square$

References

[1] Anderson, S.P., de Palma, A., Thisse, J.-F.: The CES is a discrete choice model?. Economic Letters **24**, 139-140 (1987)

[2] Anderson, S.P., de Palma, A., Thisse, J.-F.: Discrete Choice Theory of Product Differentiation. MIT Press, Cambridge, Massachusetts (1992)

[3] Berry, S.T., Levinsohn, J., Pakes, A.: Voluntary export restraints on automobiles: evaluating a strategic trade policy. American Economic Review **89**, 400–430 (1999)

[4] Dagsvik, J.K., Karlström, A.: Compensating variation and hicksian choice probabilities in random utility models that are nonlinear in income. Review of Economic Studies **72**, 57-76 (2005)

[5] de Jong, G., Daly, A., Pieters, M., van der Hoorn, T.: The logsum as an evaluation measure: Review of the literature and new results. Transportation Research Part A **41**, 874-889 (2007)

[6] de Palma, A., Kilani, K.: Switching in the logit. Economic Letters **88**, 196-202 (2005)

[7] de Palma, A., Kilani, K.: Invariance of conditional maximum utility. Journal of Economic Theory **132**, 137-146 (2007)

[8] Dixit, A.K., Stiglitz, J.E.: Monopolistic competition and optimum product diversity. American Economic Review **67**, 297-308 (1977)

[9] Domencich, T., McFadden, D.: Urban Travel Demand: A Behavioral Analysis. North-Holland, Amsterdam (1975)

[10] McFadden, D.: Modelling the choice of residential location. In: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. (eds.) Spatial Interaction Theory and Planning Models. North-Holland, Amsterdam (1978)

[11] McFadden, D.: Computing willingness-to-pay in random utility models. In: Moore, J., Riezman, R., Melvin, J. (eds.) Trade, Theory and Econometrics: Essays in the Honor of John Chipman. Routledge, London (1999)

[12] McFadden, D.: Economic choices. American Economic Review **91**, 351-378 (2001)

[13] McFadden, D.: Revealed stochastic preference: a synthesis. Economic Theory **26**, 245-264 (2005)

[14] Rudin, W.: Real and Complex Analysis. McGraw-Hill, New-York (1986)

[15] Small, K., Rosen, H.: Applied welfare analysis with discrete choice models. Econometrica **49**, 105-130 (1981)

[16] Train, K.: Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge (2003)

[17] von Haefen, R.H.: Incorporating observed choice into the construction of welfare measures from random utility models. Journal of Environmental Economics and Management **45**, 145-165 (2003)